**Leaf Diseases Classification**

## Introduction

There are many developing countries based on the agricultural sector. In India more than 60% of the population are involved in this agricultural sector. Economy highly depends on agricultural productivity hence, the farming area is apparently a very vital fragment of economy. This is one of the reasons that leaf diseases play an important role in the agricultural field. Leaf diseases in plants occur when climate changes. The build of carbon dioxide in the atmosphere is also responsible for leaf diseases.

The agricultural production at the global stage has been affected by the plant diseases the combined effect of the pests and diseases on plants can end up to loss of over 50% in the major crop, 20% in case of prime food and cash crops. This has effect on farmer income and that effects on Gross Domestic Product of the country.

The existing method to recognize any disease present in the leaf is to take the samples to the local agricultural centres. The agricultural experts visually examine the plants and classify the disease, it is totally eye based observation. For doing this, a large team of experts and continuous monitoring of plants is required which costs high, also it is a time consuming process and decision is not accurate. With the development in fields of machine learning, Artificial Intelligence and Deep Learning during the past few years identifying disease from the image of the leaf has become easy. Advancement in AI, image processing can help us to get more precise and accurate results compared to traditional method of identifying leaf diseases.

Several methods are used for classification and detection of leaf diseases from images. The most common one are logistic regression, CNN, K-Nearest Neighbours and Support Vector Machine (SVM).

## Why is Leaf diseases classification important?

There are thousands kinds of plants in ecosystems, and it is very hard to differentiate them. Only those who study plants can identify it by leaf characteristics. India provides livelihood to 6 million farmers and farmers are also facing some problems to identify the plant diseases. Plant leaf diseases classification is very important in biology research, medicine and agriculture. In medicine, herbs we have used as folk medicine. To identify herb person need practice and many years of experience, so for the ease of people and to classify the leaf. Plants are very essential in mankind as they are the source of energy supply to mankind. Plants have any diseases then it effects on the growth of plants. Using Machine learning we can overcome from this problem. Machine learning is used to automatically classify the leaf diseases types.

## Dataset:

The dataset used here contains images of leaves which are healthy as well as leaves which have got some kind of diseases. Dataset contains 15 different classes and each class contains around 1000 leaf images. Each image is of dimension 256\*256\*3 .Some of the images from dataset are shown below.







# 

## Methodology:

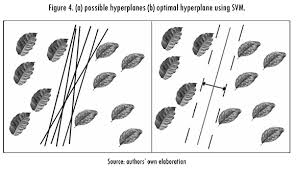
## Image pre-processing

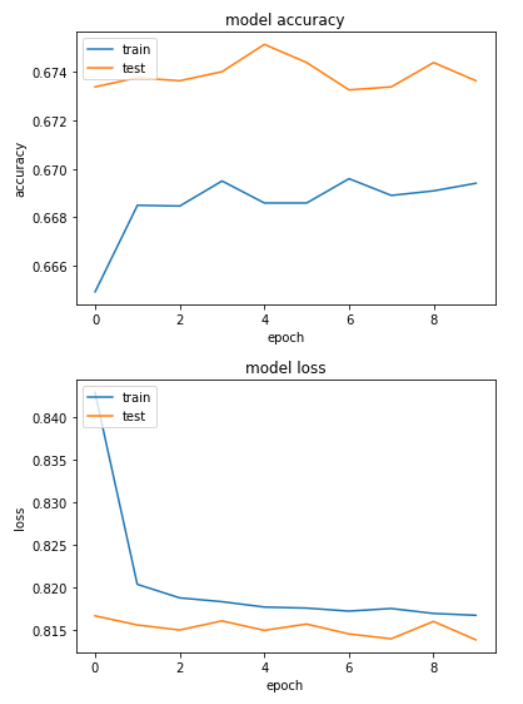
Images of the leaves can be acquired many ways such as taking photo snap or by scanning the leaf using scanner. Hence we need to pre-process all the images before using it in our model. As our model takes images in form of vector we need to convert our images into array of 256\*256\*3 size.

## Classification of leaf diseases using Support Vector Machine:

The Concept of SVM (Support Vector Machine) was introduced by Vapnik and co-workers. It gains popularity because it offers the attractive features and powerful machinery to tackle the problem of classification i.e., we need to know which belongs to which group and promising empirical performance. The SVM is based on statistical learning theory. At present SVM is popular classification tool used for pattern recognition and other classification purposes.

Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification or regression. The standard SVM classifier takes the set of input data and predicts to classify them in one of the only two distinct classes. For multiclass classification problem, we decompose multiclass problem into multiple binary class problems, and we design suitable combined multiple binary SVM classifiers.





## Classification of leaf diseases using Convolutional Neural Network:

Convolutional neural network is a class of deep neural network and most commonly applied to analysing visual imagery. The Latest generation of CNN has achieved impressive results in image classification. Using the new approach to the development of leaf image classification by the use of Convolutional neural network.

Classification of plant leaf disease using image, data and convolutional neural network.

The project contains the analysis Used to train convolutional neural network to classify different plant leaf and Diseases.

The models are trained using public dataset which have 15,000 Images of healthy and diseased leaf. The developed model is able to distinguish the healthy leaves and different plant leaf Diseases.

## Classification of leaf diseases using Artificial Neural Network:

Artificial neural network is a system closely modelled on the human brain. Artificial neural network contains the multiple layers of simple processing elements called neurone. Each neurone is linked to certain of its neighbours with coefficients of connectivity that represent the strengths of these connections. Learning is accomplished by adjusting these strengths to cause the overall network to output appropriate results. The proposed block diagram of plant classification is shows in Figure 1. This system consists mainly two phases are training and testing phases. In training phase, input plant leaves are pre-processed by resizing an image into 256x256 and apply the median filter to remove the noise. This pre-processed image is passed to feature extraction for extract the features by using haralick feature and colour coherence vector technique. Then these features are trained with ANN classifier and stored it in knowledge base. Similarly in testing phase also after extracting the features it will classifies based on the knowledge base and shows that which class is present.

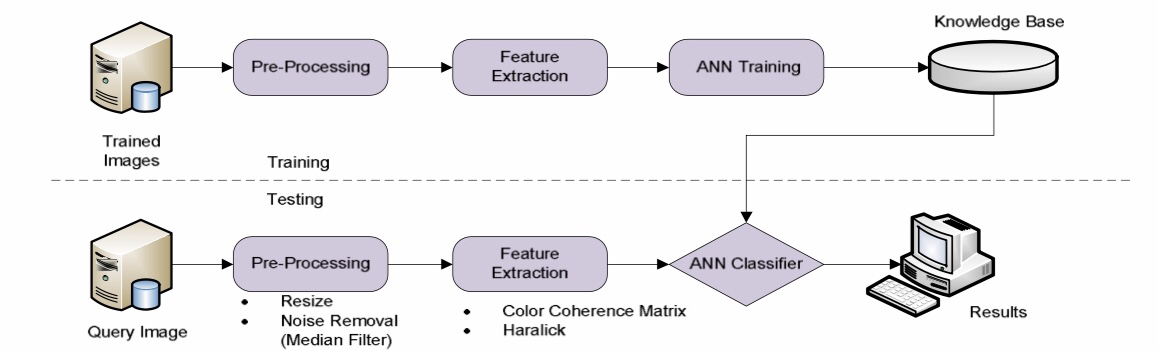
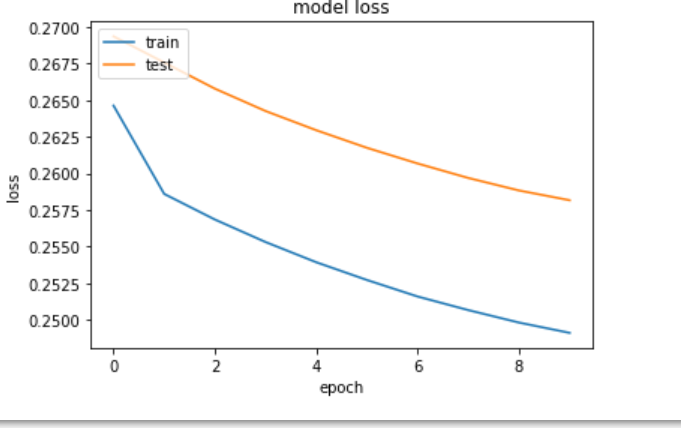
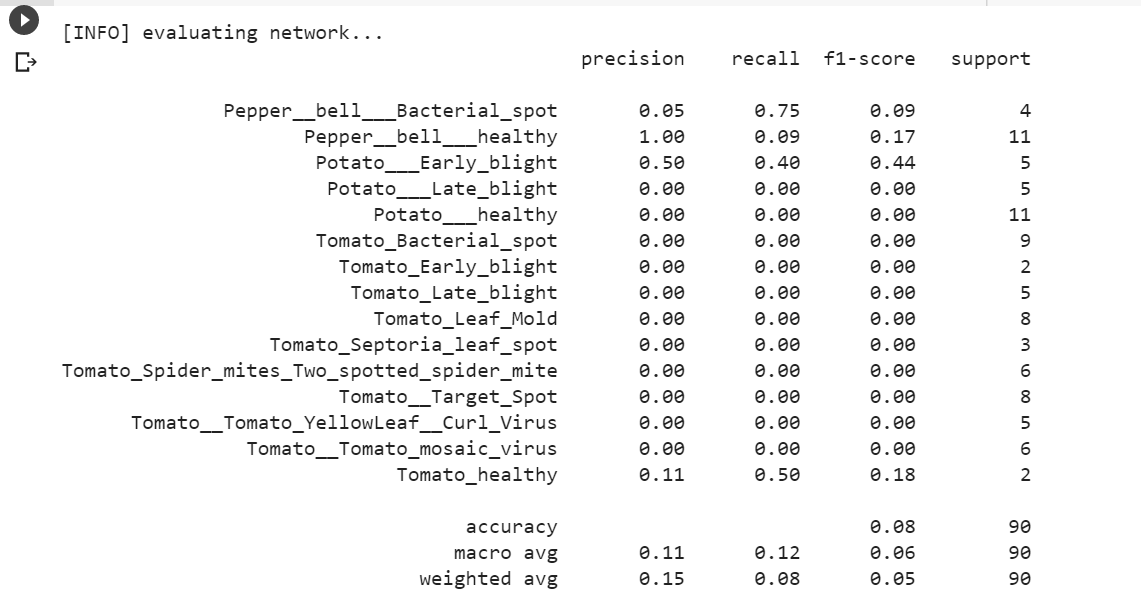


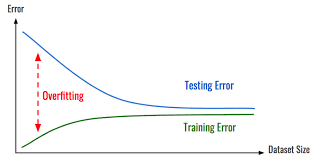
Figure 1: Proposed System





## Adding More Layers

Further we have tried adding more layers and nodes to our model to see how it effects the accuracy and overall performance of trained model. We have trained and evaluated models with different numbers of hidden layers. All the hidden layers have 32 nodes. The first model has 1 hidden layer, the second as 2 ... up to four layers. Adding more layers appears to have decreased the accuracy of the model. Giving the model greater flexibility and therefore we hope that its ability to make predictions will also increase. Unfortunately the trade-off is not that simple. Neural networks are flexible enough that they can adjust their parameters to fit the training data so precisely that they no longer generalize to data from outside the training set for example, the test data. This leads to overfitting.



## Adding Number of Nodes per Layer

To add more complexity we have added number nodes to each layer and check how our model performs compared to the previous one. So at first we started with single layer neural network, with increasing nodes in that layer. At first there were 32 nodes, then with each layer we added 64, 128, 256, 512, 1024 and 2048 nodes. We can see that as the number of nodes is increased, the model is able to better decrease the loss, e.g. to better learn the training dataset. This plot shows the direct relationship between model capacity, as defined by the number of nodes in the hidden layer and the model’s ability to learn.

## Combining both the approaches i.e. adding more nodes and more layers

After we have performed adding more layers and more nodes in isolated form, we combined both the approaches and checked the behaviour of the model. As a result we got, that till layer 5 accuracy was increasing, i.e. model was performing well with test data. But after that it test accuracy started deteriorating.

## More Training and Smaller Batches

Sometimes models with several layers need to not only train for longer, but also need more corrections per epoch. By decreasing the batch size, we can increase the number of overall corrections that a model gets. And also we need to check, it gets more fine grained information by adjusting to error over smaller batches.

In this case, we can force a model that did not learn well in previous experiments to achieve a moderately better accuracy. The performance is still not great, but it's worth mentioning that with patience and computational power we can make a model that performs more decently. Still our effort would probably be better spent on more promising models.

**Summary:**

* Neural network model capacity is controlled both by the number of nodes and the number of layers we add in our model.
* A model with a single hidden layer and a sufficient number of nodes has the capability of learning any mapping function, but the chosen learning algorithm may or may not be able to realize this capability.
* Increasing the number of layers provides a way of the increasing the capacity of the model with fewer resources and less computational resources, and modern techniques allow learning algorithms to successfully train deep models.